

Upper Limits on the Hot Jupiter Fraction in the Field of NGC 7789

D.M. Bramich^{1,2*} and Keith Horne¹

¹*School of Physics & Astronomy, University of St. Andrews, North Haugh, St. Andrews, Fife, KY16 9SS, UK*

²*Astrophysics Research Institute, Liverpool John Moores University, Twelve Quays House, Egerton Wharf, Birkenhead, CH41 1LD, UK*

Accepted 2005 August ???. Received 2005 August ???; Submitted 2005 Nov 4

ABSTRACT

We describe a method of estimating the abundance of short-period extrasolar planets based on the results of a photometric survey for planetary transits. We apply the method to a 21-night survey with the 2.5m Isaac Newton Telescope of ~ 32000 stars in a $\sim 0.5^\circ \times 0.5^\circ$ square field including the open cluster NGC 7789. From the colour-magnitude diagram we estimate the mass and radius of each star by comparison with the cluster main sequence. We search for injected synthetic transits throughout the lightcurve of each star in order to determine their recovery rate, and thus calculate the expected number of transit detections and false alarms in the survey. We take proper account of the photometric accuracy, time sampling of the observations and criteria (signal-to-noise and number of transits) adopted for transit detection. Assuming that none of the transit candidates found in the survey will be confirmed as real planets, we place conservative upper limits on the abundance of planets as a function of planet radius, orbital period and spectral type.

Key words: methods: statistical - planetary systems - open clusters and associations: individual: NGC 7789

1 INTRODUCTION

Photometric surveys for transiting extra-solar planets have become very popular since the detection of the transits exhibited by the planet-host star HD 209458 (Charbonneau et al. 2000; Brown et al. 2001). For the first time the radius of an extra-solar planet was determined, and the measurement of the orbital inclination lead to an estimate of the planetary mass, not just a lower limit. The planet HD 209458b was found to have an average density of ~ 0.38 g/cm³, significantly less than the average density of Saturn (0.7 g/cm³), leading to the term “hot Jupiter” for the class of Jupiter mass planets with short periods (1-10 d). Transiting planets are also very important in that their atmospheric composition may be determined from transmission spectroscopy for the brighter host stars (Charbonneau et al. 2002; Brown, Libbrecht & Charbonneau 2003; Vidal-Madjar et al. 2003). Careful modelling of the transit morphology and/or timings may be used to constrain the presence of moons or rings and to probe the limb darkening of the star (Brown et al. 2001).

Since the discovery of the transiting nature of HD 209458b, many transit candidates have been put forwards by various groups (e.g: Street et al. 2003; Drake & Cook 2004; Bramich et al. 2005). OGLE have been by far the most prolific transit survey with 177 transit candidates from three observational seasons (Udalski et al. 2002a; Udalski et al. 2002b; Udalski et al. 2003; Udalski et al. 2004). However, even with the discovery of numerous candi-

dates, follow-up observations have confirmed the planetary status of only six, bringing the total number of transiting planets to nine (see Bramich et al. 2005 and references therein; Sato et al. 2005; Bouchy et al. 2005). This is due to the ubiquity of eclipsing binaries and the many observational scenarios involving these systems that mimic a transit event (Brown 2003). Spectroscopic and multi-band photometric observations are required to rule out the eclipsing binary scenarios and determine the mass of the companion (e.g: Alonso et al. 2004).

When hunting for new planets, the main advantage of the transit method over the radial-velocity technique is that many stars may be monitored in parallel and to fainter magnitudes, thus probing out beyond the Solar neighbourhood. Even though only a small fraction ($\sim 0.1\%$) of stars are expected to exhibit a hot Jupiter transit signal, by using a large field of view instrument on a crowded star field one can monitor enough stars to the precision required to detect a number of transiting planets. Consequently large charge-coupled device (CCD) mosaic cameras are essential to the planet catch potential of a transit survey.

A transit survey produces transit candidates that need follow-up observations to determine the nature of the transit signals. Candidates confirmed as transiting planets add to our database of extra-solar planets and constrain their poorly known mass-radius relationship (Burrows et al. 2004). To estimate the fraction of stars that harbour a planet (the planet fraction) as a function of spectral type and planet type we compare the number of transiting planets detected with a calculation of the expected number of transiting planet

* E-mail: dmb7@st-and.ac.uk

Table 1. The different subsets of stars used when calculating the expected number of transiting planet detections and false alarms.

Set Of Stars	Mass Range	No. Of Stars	No. Of Stars With 1999-07 Lightcurve Data
All Stars	$0.08M_{\odot} \leq M_* \leq 1.40M_{\odot}$	32027	20949
Late F Stars	$1.05M_{\odot} \leq M_* \leq 1.40M_{\odot}$	3129	2780
G Stars	$0.80M_{\odot} \leq M_* \leq 1.05M_{\odot}$	7423	6711
K Stars	$0.50M_{\odot} \leq M_* \leq 0.80M_{\odot}$	15381	9690
M Stars	$0.08M_{\odot} \leq M_* \leq 0.50M_{\odot}$	6094	1768

detections. Even when zero planets are detected (a null result) such a calculation places upper limits on the planet fraction.

In this paper we describe a Monte Carlo method for calculating detection probabilities (and false alarm rates) of transiting planets based on photometric data, as functions of various parameters, taking into account the following factors:

- (i) Limb darkening effects which tend to make central eclipses deeper and grazing eclipses shallower.
- (ii) The effect of orbital inclination on the shape and width of the transit lightcurve.
- (iii) The distribution of the photometric data in time and the individual error bars on each measurement.
- (iv) The signal-to-noise threshold, number of transits, and number of data points in-transit and out-of-transit required for a detection.

We then apply the method to the transit survey described in Bramich et al. 2005 (here on referred to as BRA05) to determine the expected number of transiting planet detections and place limits on the planet fraction as a function of star and planet type.

In Section 2 we describe the lightcurve data used in the analysis and in Section 3 we define the detection probabilities and false alarm probabilities for an extra-solar planet based on photometric data. In Section 4 we present the Monte Carlo method that we used to calculate these probabilities and derive limits on the hot Jupiter fraction in the field of NGC 7789 as a function of star and planet type. In Section 5 we discuss the results and in Section 6 we present our conclusions.

2 THE TRANSIT SURVEY DATA ON NGC 7789

A transit survey of the field of NGC 7789 was presented in BRA05 in which ~ 33000 stars were photometrically monitored in the Sloan r' band over three separate runs with dates 1999 June 22-30, 1999 July 22-31 and 2000 September 10-20. For brevity these runs shall be referred to from now on as 1999-06, 1999-07 and 2000-09 respectively.

To summarise, in BRA05, Sloan $r' - i'$ colour indices were used to construct a colour-magnitude diagram (CMD) and thereby identify the cluster main sequence. Fig. 1 shows the CMD for the stars from chip 4 which was centred on the cluster. Although the cluster main sequence is visible, it is clear that most stars in the sample are field stars and not cluster stars. A theoretical main sequence model for the stellar mass range $0.08 M_{\odot} \leq M_* \leq 1.40 M_{\odot}$ was adopted and fitted to the cluster main sequence via magnitude offsets. Using the known cluster distance $d_c = 2337\text{pc}$ and reddening $E(B - V) = 0.217$, and adopting an Einasto law for the distribution of the interstellar medium in the Milky Way (Robin et al. 2003), a distance d_* was determined for each star such that the theoretical main sequence passes through the star's position on the CMD. It was argued that

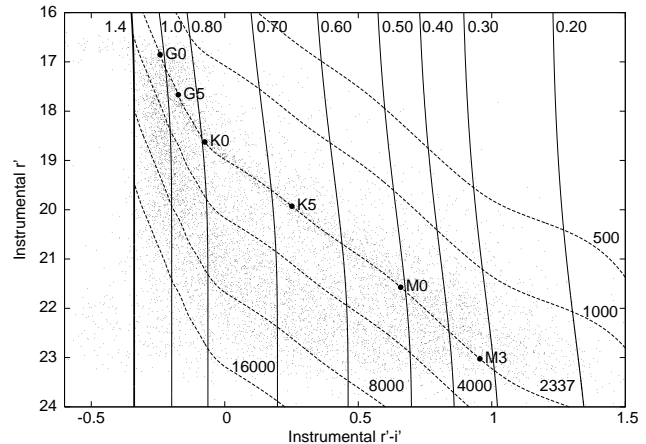


Figure 1. Instrumental CMD for chip 4 (taken from BRA05) which was centred on the cluster. The numbers along the top are masses in units of M_{\odot} and the numbers along the bottom and right are distances in parsecs. Fiducial spectral types are marked on the cluster theoretical main sequence for clarity.

giant stars lie beyond the edge of the galaxy in order to be non-saturated in the image data. Hence, it was assumed that each star is on the main sequence, and after determining the star's distance d_* , the star's mass M_* and radius R_* could be read off from its position on the theoretical main sequence.

In this paper we consider the 32027 stars from this data set that have a lightcurve from the 2000-09 run, and an assigned distance, mass and radius. The remaining stars with lightcurves lack a colour measurement or were too blue to be assigned a mass and radius using the adopted theoretical main sequence. We also consider the lightcurve data from the 1999-07 run where it exists. BRA05 searched for transits in the 10-night 1999-07 run and the 11-night 2000-09 run, 14 months later. The 1999-06 run was too sparsely sampled in time to support transit hunting by the adopted search technique.

We are interested in the expected number of transiting planet detections (and false alarms) for stars of different masses or, equivalently, spectral types. To facilitate this analysis we consider 4 mutually exclusive subsets of stars which in union make up the set of all 32027 stars. These sets are the late F stars, G stars, K stars and M stars respectively. Table 1 shows the number of stars in each set and the spectral type/mass ranges to which they correspond. The table also includes the number of stars for which 1999-07 lightcurve data exists. The mass ranges for the various spectral types are taken from Lang (1992).

3 DETECTION PROBABILITIES AND FALSE ALARM RATES

In BRA05, a matched filter algorithm was used to search for transits in the lightcurves by adopting a square “boxcar” shape for the transit lightcurve of total width $5\Delta t$ (where Δt is the transit duration searched for). This search was based on the transit detection statistic:

$$S_{\text{tra}}^2 \equiv \frac{\chi_{\text{const}}^2 - \chi_{\text{tra}}^2}{\left(\frac{\chi_{\text{out}}^2}{N_{\text{out}} - 1}\right)} \quad (1)$$

where χ_{tra}^2 is the chi squared of the boxcar transit fit, χ_{const}^2 is the chi squared of the constant fit, χ_{out}^2 is the chi squared of the boxcar transit fit for the N_{out} out-of-transit data points. Transit candidates were chosen using a threshold of $S_{\text{tra}} \geq S_{\text{min}} = 10$.

Consider an extra-solar planet of radius R_p , orbital period P and orbital inclination i with t_0 as the time of mid-transit. The planet orbits a star S, of known mass M_* and radius R_* , that has an associated lightcurve. We calculate the predicted transit lightcurves based on a simple planet-star model: we assume a luminous primary, linear limb darkening with $u = 0.5$ and a dark massless companion in a circular orbit. Adding this signal into the observed lightcurve of the star, we calculate the transit statistic S_{tra} (Eqn. 1) for each transit event, and then evaluate the following detection function:

$$D(S, S_{\text{min}}, N_{\text{min}}, R_p, P, i, t_0) = \begin{cases} 1 & \text{if } S_{\text{tra}} \geq S_{\text{min}} \text{ for at least } N_{\text{min}} \text{ predicted transits} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Using the same procedure as above, but without actually adding the predicted transit lightcurve into the observed lightcurve of the star, we evaluate the false alarm function:

$$F(S, S_{\text{min}}, N_{\text{min}}, R_p, P, i, t_0) = \begin{cases} 1 & \text{if } S_{\text{tra}} \geq S_{\text{min}} \text{ for at least } N_{\text{min}} \text{ predicted transits} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

The function D is a trigger function that indicates where the data and detection algorithm are capable of detecting a transit of the specified type, and F indicates where the data alone suggest that such a transit is actually present. The BRA05 lightcurve data contains some eclipsing binary stars and possibly transits. Hence both D and F are slightly over estimated.

In the upper panel of Fig. 2 we plot a subsection of the lightcurve of star 61377 with an injected 0.02 mag offset of duration 3 h starting at $t_0 = 2451799.5$ d. This $r' \approx 18.20$ mag G star has a mass, radius and distance of $0.96M_{\odot}$, $0.96R_{\odot}$ and 3152pc respectively. In the lower panel of Fig. 2 we plot the corresponding periodic functions $D(t_0)$ and $F(t_0)$ represented by thick and thin continuous lines respectively. We adopted $S_{\text{min}} = 10$, $N_{\text{min}} = 1$, $R_p = 1.2R_J$, $P = 3.338$ d and $i = 90^\circ$ for this calculation. The function $D(t_0)$ attains the value of 1 where there is data of sufficient precision to detect a transit. The function $F(t_0)$ attains the value of 1 where there is data that mimics a transit signature, and it has clearly been triggered by the injected offset in the lightcurve data.

In Eqn. 2, we integrate out the “nuisance parameters” t_0 and i

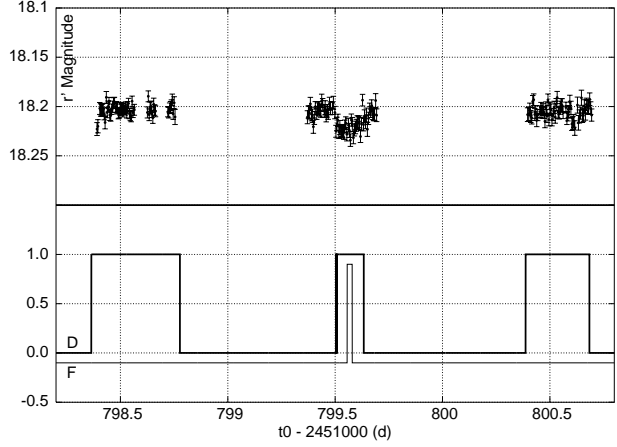


Figure 2. Upper panel: A subsection of the lightcurve of star 61377 with an injected 0.02 mag offset of duration 3 h starting at $t_0 = 2451799.5$. **Lower panel:** The periodic functions $D(t_0)$ and $F(t_0)$ for star 61377 represented by thick and thin continuous lines respectively. $F(t_0)$ is offset vertically from $D(t_0)$ by 0.1 for clarity. We adopted $S_{\text{min}} = 10$, $N_{\text{min}} = 1$, $R_p = 1.2R_J$, $P = 3.338$ d and $i = 90^\circ$ for this calculation.

to get:

$$P(\text{det} | S, S_{\text{min}}, N_{\text{min}}, R_p, P) = \int_0^{90^\circ} di \int_0^P dt_0 f(t_0, i) D(S, S_{\text{min}}, N_{\text{min}}, R_p, P, i, t_0) \quad (4)$$

where $P(\text{det} | S, S_{\text{min}}, N_{\text{min}}, R_p, P)$ is the detection probability for star S and $f(t_0, i)$ is the joint probability distribution function (PDF) of t_0 and i . We assume that the parameters t_0 and i are independent, t_0 is uniformly distributed over $0 \leq t_0 \leq P$ and random orbit orientation for i in the range $0^\circ \leq i \leq 90^\circ$. Hence we may write:

$$f(t_0, i) = \frac{\pi \sin i}{180^\circ P} \quad (5)$$

Combining Eqns. 4 and 5 we get the detection probability:

$$P(\text{det} | S, S_{\text{min}}, N_{\text{min}}, R_p, P) = \int_0^{90^\circ} di \int_0^P dt_0 \left(\frac{\pi \sin i}{180^\circ P} \right) D(S, S_{\text{min}}, N_{\text{min}}, R_p, P, i, t_0) \quad (6)$$

Using a parallel argument, we obtain an expression for the false alarm probability $P(\text{fal} | S, S_{\text{min}}, N_{\text{min}}, R_p, P)$ as:

$$P(\text{fal} | S, S_{\text{min}}, N_{\text{min}}, R_p, P) = \int_0^{90^\circ} di \int_0^P dt_0 \left(\frac{\pi \sin i}{180^\circ P} \right) F(S, S_{\text{min}}, N_{\text{min}}, R_p, P, i, t_0) \quad (7)$$

4 MONTE CARLO SIMULATIONS

4.1 Methodology

We take the Monte Carlo approach to evaluating the detection probabilities and false alarm rates (Press et al. 1992), rather than attempting to numerically integrate Eqns. 6 and 7. In general, a Monte Carlo simulation estimates the probability of an event by

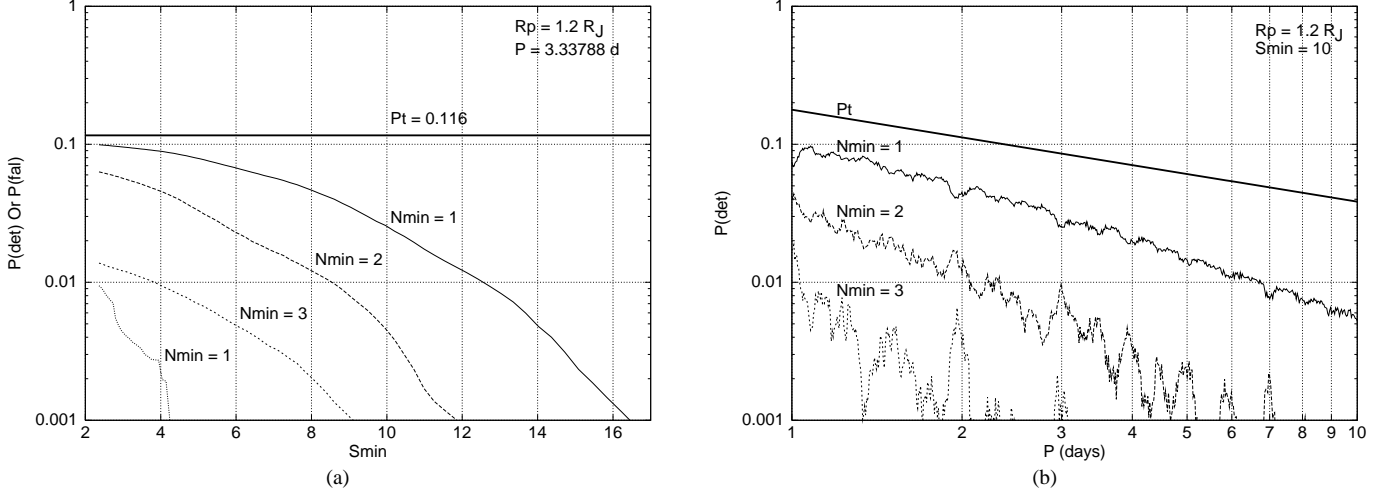


Figure 3. (a): Detection probability as a function of S_{\min} for star 61377 with $P = 3.338$ d and $R_p = 1.2 R_J$. The continuous, dashed and shorter dashed lines correspond to $N_{\min} = 1$, $N_{\min} = 2$ and $N_{\min} = 3$ respectively. The false alarm probability for the same star and planet with $N_{\min} = 1$ is shown by the dotted line. The thick continuous line corresponds to the probability $P_t = 0.116$ that the planet-star system exhibits transits. (b): Detection probability as a function of P for star 61377 with $S_{\min} = 10$ and $R_p = 1.2 R_J$. Again, the continuous, dashed and shorter dashed lines correspond to $N_{\min} = 1$, $N_{\min} = 2$ and $N_{\min} = 3$ respectively. The false alarm probability is approximately zero for all P and N_{\min} at this detection threshold. The thick continuous curve corresponds to the probability P_t that the planet-star system exhibits transits.

selecting a large random sample from the parameter space as governed by the underlying PDF, and then calculating the fraction of the sample that satisfy the event criteria. The larger the sample, the more accurate the calculated probability. However, the size of the sample that may be selected and analysed is usually limited by available computing resources.

For each star S and its corresponding lightcurve, we used the Monte Carlo method to calculate $P(\text{det} | S, S_{\min}, N_{\min}, R_p, P)$ and $P(\text{fal} | S, S_{\min}, N_{\min}, R_p, P)$ for a grid in S_{\min} , N_{\min} , R_p and P . We chose to use a geometric sequence in S_{\min} from $S_{\min} = 2.4$ to $S_{\min} = 26.4$ with geometric factor 1.025. We also chose a geometric sequence in P from $P = 1$ d to $P = 10$ d with geometric factor 1.004 resulting in 576 period values. We chose $N_{\min} \in \{1, 2, 3\}$ and $R_p = 1.2 R_J$. For each grid point we selected a set of $N_{\text{MC}} = 1000$ planets with t_0 and i drawn randomly for each planet from the PDF in Eqn. 5.

The grid for P should be fine enough that the difference in period $\Delta P = P_{i+1} - P_i$ between two consecutive grid points P_i and P_{i+1} (where $P_{i+1} > P_i$) is such that the difference in the number of cycles spanning the duration of the lightcurve is less than or equal to a fraction f_t of the transit duration (in cycle units). This condition implies that the grid in P should be a geometric sequence with geometric factor less than or equal to $1 + (f_t \Delta t / T)$ where Δt is the transit duration and T is the duration of the lightcurve. Adopting $f_t = 0.5$, $\Delta t \approx 2$ h for a typical transit duration and $T = 10.4$ d corresponding to the longer 2000-09 run yields $f_t \Delta t / T \approx 4.0 \times 10^{-3}$. Hence our choice of grid in P is fine enough for the $\sim 35\%$ of stars that have lightcurve data from the 2000-09 run alone. For the remaining stars with lightcurve data from both runs, adopting the much finer period grid that is required makes the Monte Carlo simulations prohibitive in terms of computer processing time. To fully sample the possible period aliases introduced by using the adopted grid, the current period value P_{curr} for each Monte Carlo trial was drawn from a uniform distribution on the interval $P / \sqrt{1.004} \leq P_{\text{curr}} \leq P \sqrt{1.004}$.

For the 1-10 d planets that we consider, the probability that a

planet transits is $\sim 10\%$. Hence, the number of Monte Carlo trials that result in a detection is smaller than N_{MC} by at least a factor of 10, leading to a relatively noisy determination of the detection probabilities (for $N_{\text{MC}} = 1000$, $\sigma \sim 10\%$). However, when we sum the detection probabilities over a large number of stars, typically $\sim 10^4$ (see Section 4.3), the noise is reduced to an insignificant level ($\sim 0.1\%$). Similarly, it is reduced even further when we integrate over various period ranges (see Section 4.4).

4.2 An Example Simulation

Let us now consider the complete lightcurve of star 61377. The star has 118 data points over 10 nights in its 1999-07 lightcurve with a standard deviation of ~ 0.007 mag and 612 data points over 11 nights in its 2000-09 lightcurve with a standard deviation of ~ 0.010 mag.

In Fig. 3, we plot the detection probability and false alarm probability as functions of the transit statistic detection threshold (Fig. 3(a)) and period (Fig. 3(b)) for star 61377. We used $N_{\text{MC}} = 10^5$ Monte-Carlo trials for this star in order to reduce the noise from the simulations for illustrative purposes. The detection and false alarm probabilities both decrease with detection threshold S_{\min} (Fig. 3(a)). For this particular star, it can be seen that false alarms are very unlikely even at very low detection thresholds. The thick continuous line corresponds to the probability $P_t = 0.116$ that the planet-star system exhibits transits, calculated from:

$$P_t = \frac{R_p + R_*$$

$$= 0.162 \left(\frac{R_p + R_*}{R_\odot} \right) \left(\frac{M_*}{M_\odot} \right)^{-1/3} \left(\frac{P}{1 \text{ d}} \right)^{-2/3} \quad (8)$$

where a is the orbital radius and the final expression uses Kepler's law. The completeness of the transit search falls rapidly with the

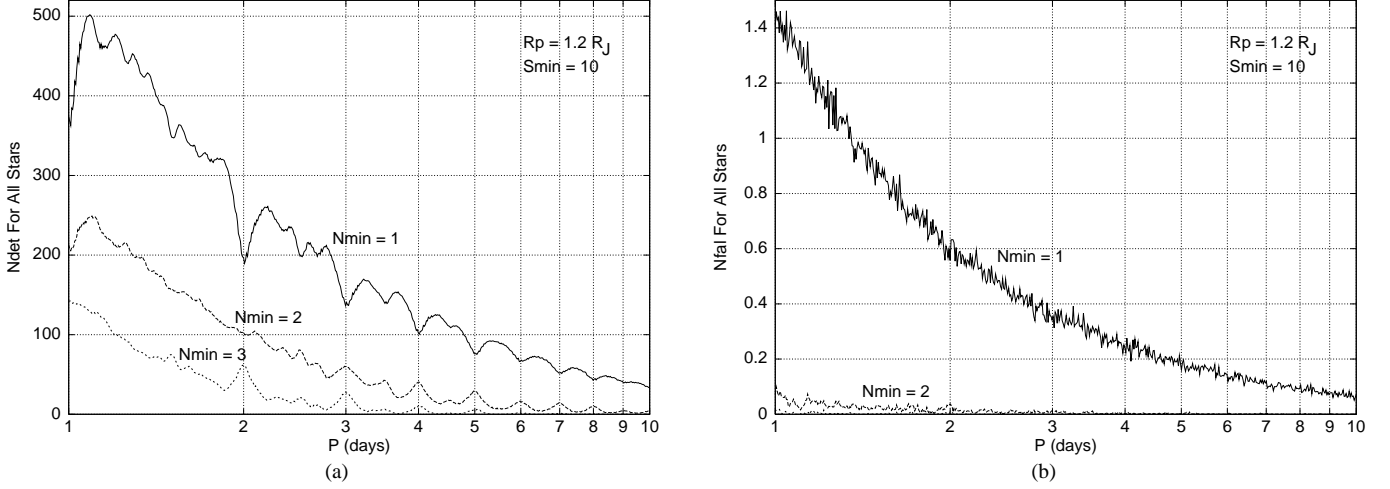


Figure 4. (a): Expected number of transiting planet detections for all stars as a function of P for $S_{\text{min}} = 10$ and $R_p = 1.2 R_J$. The continuous, dashed and shorter dashed lines correspond to $N_{\text{min}} = 1$, $N_{\text{min}} = 2$ and $N_{\text{min}} = 3$ respectively. (b): Expected number of false alarms for all stars as a function of P for $S_{\text{min}} = 10$ and $R_p = 1.2 R_J$. Again, the continuous, dashed and shorter dashed lines correspond to $N_{\text{min}} = 1$, $N_{\text{min}} = 2$ and $N_{\text{min}} = 3$ respectively.

adopted detection threshold S_{min} . For $N_{\text{min}} = 1$, the completeness is $\sim 76.7\%$ at threshold $S_{\text{min}} = 4$, dropping to $\sim 22.3\%$ for $S_{\text{min}} = 10$.

In Fig. 3(b) we see the expected $P^{-2/3}$ dependence of detection probability on orbital period, but with more detailed period structure arising from the detailed time sampling of the observations. For $N_{\text{min}} = 1$, orbital periods close to integer values tend to have lower detection probabilities since such periods are resonant with the observational gaps during the daytime. Conversely, orbital periods close to fractional values tend to have higher detection probabilities since such periods cover a greater range of orbital phases. For example, periods close to ~ 3.0 d have a detection probability of ~ 0.025 whereas periods close to ~ 2.7 d have a detection probability of ~ 0.036 for this particular star. Fig. 3(b) also shows that as you increase the number of recovered transits required for a detection, the detection probability decreases rapidly.

4.3 Expected Number Of Transiting Planet Detections

Assuming that each star S has one planet of radius R_p and period P , then the expected number of transiting planet detections $N_{\text{det}}(Y, S_{\text{min}}, N_{\text{min}}, R_p, P)$ as a function of star type Y , S_{min} , N_{min} , R_p and P is simply the sum of the detection probabilities for all stars of the required type:

$$N_{\text{det}}(Y, S_{\text{min}}, N_{\text{min}}, R_p, P) = \sum_{S \in Y} P(\text{det} | S, S_{\text{min}}, N_{\text{min}}, R_p, P) \quad (9)$$

Similarly, the expected number of false alarms $N_{\text{fal}}(Y, S_{\text{min}}, N_{\text{min}}, R_p, P)$ is given by:

$$N_{\text{fal}}(Y, S_{\text{min}}, N_{\text{min}}, R_p, P) = \sum_{S \in Y} P(\text{fal} | S, S_{\text{min}}, N_{\text{min}}, R_p, P) \quad (10)$$

In Fig. 4, we plot the expected number of transiting planet detections (Fig. 4(a)) and the expected number of false alarms (Fig. 4(b)) for all stars as functions of the period with $S_{\text{min}} = 10$ and

$R_p = 1.2 R_J$. These quantities have a strong dependence on period in the same way as the probabilities from which they are derived (see Section 4.2). Fig. 4(b) clearly shows that, by increasing N_{min} from 1 to 2, the expected number of false alarms is effectively reduced to zero for all P . However, introducing this extra constraint for a transit detection reduces the expected yield of planets from the survey by more than a factor of 2 (Fig. 4(a)). Also, setting $N_{\text{min}} = 2$ is unnecessary since the expected number of false alarms for the detection threshold chosen in BRA05 ($S_{\text{min}} = 10$ and $N_{\text{min}} = 1$) is less than 1 for $P > 1.34$ d, indicating a good choice of detection threshold for all except the shortest period hot Jupiters.

4.4 Expected Number Of Transiting Hot Jupiter Detections

We can now estimate the expected number of detections and false alarms for three different planet period ranges, and for different star types. We consider the very hot Jupiters with periods of 1-3 d, the shorter period hot Jupiters with periods of 3-5 d and the longer period hot Jupiters with periods of 5-10 d. Within each period range $P_1 \leq P \leq P_M$ we assume that planets are uniformly distributed in $\log(P)$. This is roughly consistent with the results of the radial velocity surveys (Heacox 1999; Bramich 2005).

Assuming that each star S has one planet of radius R_p in the specified period range, then the expected number of transiting planet detections is obtained by summing over star type Y and integrating over period P :

$$N_{\text{det}}(Y, S_{\text{min}}, N_{\text{min}}, R_p, P_1, P_M) = \sum_{S \in Y} \int_{P_1}^{P_M} \left(\frac{d \ln P}{\ln(P_M/P_1)} \right) P(\text{det} | S, S_{\text{min}}, N_{\text{min}}, R_p, P) \quad (11)$$

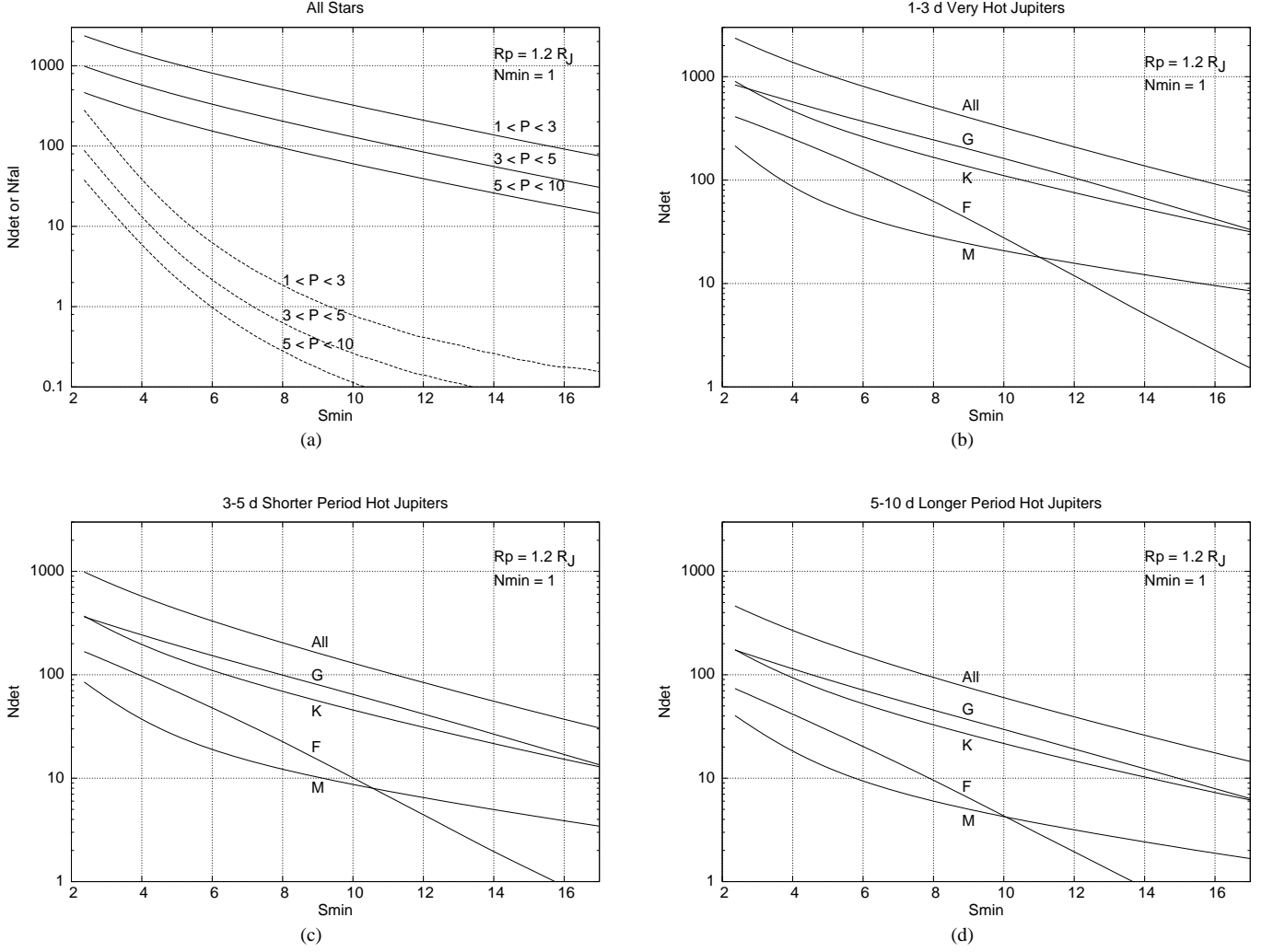


Figure 5. (a): Expected number of transiting hot Jupiter detections (continuous curves) and false alarms (dashed curves) for all stars as functions of S_{\min} with $N_{\min} = 1$ and $R_p = 1.2 R_J$. Each curve is labelled with the hot Jupiter period range to which it corresponds. (b): Expected number of transiting 1-3 d hot Jupiter detections as a function of S_{\min} with $N_{\min} = 1$ and $R_p = 1.2 R_J$ for various sets of stars. Each curve is labelled with the star type to which it corresponds. (c): Expected number of transiting 3-5 d hot Jupiter detections as a function of S_{\min} with $N_{\min} = 1$ and $R_p = 1.2 R_J$ for various sets of stars. Each curve is labelled with the star type to which it corresponds. (d): Expected number of transiting 5-10 d hot Jupiter detections as a function of S_{\min} with $N_{\min} = 1$ and $R_p = 1.2 R_J$ for various sets of stars. Each curve is labelled with the star type to which it corresponds.

Similarly, the expected number of false alarms is given by:

$$\begin{aligned}
 N_{\text{fal}}(Y, S_{\min}, N_{\min}, R_p, P_1, P_M) \\
 = \sum_{S \in Y} \int_{P_1}^{P_M} \left(\frac{d \ln P}{\ln(P_M/P_1)} \right) P(\text{fal} | S, S_{\min}, N_{\min}, R_p, P)
 \end{aligned}
 \quad (12)$$

In Fig. 5(a) we plot the expected number of transiting planet detections (continuous curves) and false alarms (dashed curves) for all stars as functions of S_{\min} with $N_{\min} = 1$ and $R_p = 1.2 R_J$. Each curve is labelled with the period range to which it corresponds. Similarly, each of Figs. 5(b) - (d) corresponds to a different period range in which we plot the expected number of transiting planet detections for the F, G, K and M stars in our sample as a function of S_{\min} with $N_{\min} = 1$ and $R_p = 1.2 R_J$.

For a limited range $7 \leq S_{\min} \leq 15$, the “curves” for N_{det} in all of Figs. 5(a)-(d) can be approximated by straight lines, allowing

us to express our results in the form of an approximate empirical relationship:

$$\frac{N_{\text{det}}(S_{\min})}{N_{\text{det}}(10)} \approx \exp \left[A \left(1 - \frac{S_{\min}}{10} \right) \right] \quad (13)$$

where A is a constant that may be determined for each set of stars Y and period range P_1 to P_M considered. We note that the transit detection statistic S_{tra} is equivalent to the signal-to-noise (S/N) of the fitted transit signal and that:

$$S_{\text{tra}} \equiv \text{S/N} \propto \frac{\Delta f}{f_0} \sim \left(\frac{R_p}{R_*} \right)^2 \quad (14)$$

where $\Delta f/f_0$ is the fractional transit depth. Now, since we have $S_{\text{tra}} \propto R_p^2$, we may infer that for a fixed S_{tra} the equivalent detection threshold varies as $S_{\min} \propto R_p^{-2}$. Using this fact, we

Table 2. Results of the Monte Carlo simulations for $S_{\min} = 10$, $N_{\min} = 1$ and $R_p = 1.2R_J$. The number in brackets for N_{\det} and N_{fal} indicates the uncertainty on the last decimal place. The upper limits on f_p at the significance level α may be scaled with S_{\min} and R_p by using the scaling relation defined in Eqn. 21 together with the listed values of A .

Y	No. Of Stars	Period Range	N_{\det}	N_{fal}	A	Upper Limit On f_p At $\alpha = 0.50$	Upper Limit On f_p At $\alpha = 0.05$	Upper Limit On f_p At $\alpha = 0.01$
All Stars	32027	$1\text{d} \leq P \leq 3\text{d}$	325.09(2)	0.792(1)	2.18	0.213%	0.922%	1.42%
All Stars	32027	$3\text{d} \leq P \leq 5\text{d}$	130.77(2)	0.267(1)	2.17	0.530%	2.29%	3.52%
All Stars	32027	$5\text{d} \leq P \leq 10\text{d}$	60.59(1)	0.115(1)	2.15	1.14%	4.94%	7.60%
Late F Stars	3129	$1\text{d} \leq P \leq 3\text{d}$	28.50(1)	0.223(1)	4.20	2.43%	10.5%	16.2%
Late F Stars	3129	$3\text{d} \leq P \leq 5\text{d}$	10.37(1)	0.075(1)	4.10	6.68%	28.9%	44.4%
Late F Stars	3129	$5\text{d} \leq P \leq 10\text{d}$	4.38(1)	0.028(1)	3.99	15.8%	68.4%	105%
G Stars	7423	$1\text{d} \leq P \leq 3\text{d}$	163.57(2)	0.294(1)	2.18	0.424%	1.83%	2.82%
G Stars	7423	$3\text{d} \leq P \leq 5\text{d}$	65.37(1)	0.110(1)	2.20	1.06%	4.58%	7.04%
G Stars	7423	$5\text{d} \leq P \leq 10\text{d}$	29.85(1)	0.051(1)	2.19	2.32%	10.0%	15.4%
K Stars	15381	$1\text{d} \leq P \leq 3\text{d}$	111.76(1)	0.265(1)	1.91	0.620%	2.68%	4.12%
K Stars	15381	$3\text{d} \leq P \leq 5\text{d}$	46.15(1)	0.078(1)	1.94	1.50%	6.49%	9.98%
K Stars	15381	$5\text{d} \leq P \leq 10\text{d}$	21.80(1)	0.035(1)	1.94	3.18%	13.7%	21.1%
M Stars	6094	$1\text{d} \leq P \leq 3\text{d}$	21.26(1)	0.010(1)	1.44	3.26%	14.1%	21.7%
M Stars	6094	$3\text{d} \leq P \leq 5\text{d}$	8.88(1)	0.004(1)	1.51	7.81%	33.7%	51.9%
M Stars	6094	$5\text{d} \leq P \leq 10\text{d}$	4.28(1)	0.002(1)	1.51	16.2%	70.0%	108%

may rewrite Eqn. 13 as:

$$\frac{N_{\det}(S_{\min}, R_p)}{N_{\det}(10, 1.2R_J)} \approx \exp \left[A \left(1 - \left(\frac{S_{\min}}{10} \right) \left(\frac{1.2R_J}{R_p} \right)^2 \right) \right] \quad (15)$$

For $S_{\min} = 10$, the range of planetary radii over which this scaling relation is valid is $0.98 R_J \leq R_p \leq 1.43 R_J$.

The relation in Eqn. 15 allows us to scale our results for a range of planetary radii that encompasses the radii of the known transiting hot Jupiters without having to repeat the Monte Carlo simulations. In Table 2 we present the results of the Monte Carlo simulations for the various sets of stars and orbital period ranges considered, as calculated for $S_{\min} = 10$, $N_{\min} = 1$ and $R_p = 1.2 R_J$. We also include the value of A from Eqn. 15 for each combination of star and planet type. One can see that we expect to detect ~ 325 1-3 d very hot Jupiters, ~ 131 3-5 d hot Jupiters and ~ 61 5-10 d hot Jupiters from the transit survey based on the assumption that each star has a single planet of the specified type.

4.5 Upper Limits On The Hot Jupiter Fraction

In reality, only a fraction f_p of the stars considered in our Monte Carlo simulations harbour a planet of a specified type, instead of our assumed one planet per star. Hence we must correct our calculations of the expected number of transiting planet detections by this factor. We refer to f_p as the planet fraction. Since f_p is an unknown quantity that we would like to estimate, we may use the fact that the transit survey in BRA05 has most likely produced a null result (although this is still to be confirmed) and place a significant upper limit on f_p .

First of all we make the assumption that the actual number of transiting planet detections X has a Poisson distribution with expected value $E(X)$ given by:

$$E(X) = f_p N_{\det} \quad (16)$$

The Poisson distribution is defined by:

$$P(X = x) = \frac{(E(X))^x}{x!} e^{-E(X)} \quad \text{for } x \in \mathbb{N}_0 \quad (17)$$

For a null result, $x = 0$. Using this fact and substituting Eqn. 16 into Eqn. 17 we get:

$$P(X = 0) = e^{-f_p N_{\det}} \quad (18)$$

In order to obtain an upper limit on f_p at the significance level α we require:

$$P(X = 0) \leq \alpha \quad (19)$$

Hence, from Eqns. 18 and 19, we derive:

$$f_p \leq \frac{-\ln \alpha}{N_{\det}} \quad (20)$$

The values of f_p that we derive in this manner for $\alpha = 0.50$, $\alpha = 0.05$ and $\alpha = 0.01$ are shown in Table 2. For example, for $\alpha = 0.01$, we place an upper limit of 1.42% on the 1-3 d very hot Jupiter fraction based on the assumption that such planets have a typical radius of $1.2 R_J$. Our most conservative upper limits are obtained for $\alpha = 0.01$ and consequently these are the upper limits on f_p that we consider in the discussion and conclusions.

Finally, we derive how the upper limit on f_p scales with S_{\min} and R_p by using Eqn. 15 in Eqn. 20:

$$f_p \leq \frac{-\ln \alpha}{N_{\det}(10, 1.2R_J)} \exp \left[A \left(\left(\frac{S_{\min}}{10} \right) \left(\frac{1.2R_J}{R_p} \right)^2 - 1 \right) \right] \quad (21)$$

5 DISCUSSION

We have derived relatively stringent upper limits on the abundance of hot Jupiters for the field of NGC 7789. The most stringent upper limit on f_p at $\alpha = 0.01$ of 0.79% is obtained for 1-3 d very hot Jupiters of radius $1.4 R_J$. For the Sun-like G stars, we obtain limits on f_p at $\alpha = 0.01$ of 2.82% for the 1-3 d very hot Jupiters and 7.04% for the 3-5 d hot Jupiters (assuming $R_p = 1.2R_J$). Fig. 6 shows plots of the upper limit on the planet fraction f_p at a significance level of 1% as a function of star type (Fig. 6(a)) and orbital period (Fig. 6(b)) for an assumed planetary radius of $1.2 R_J$. We

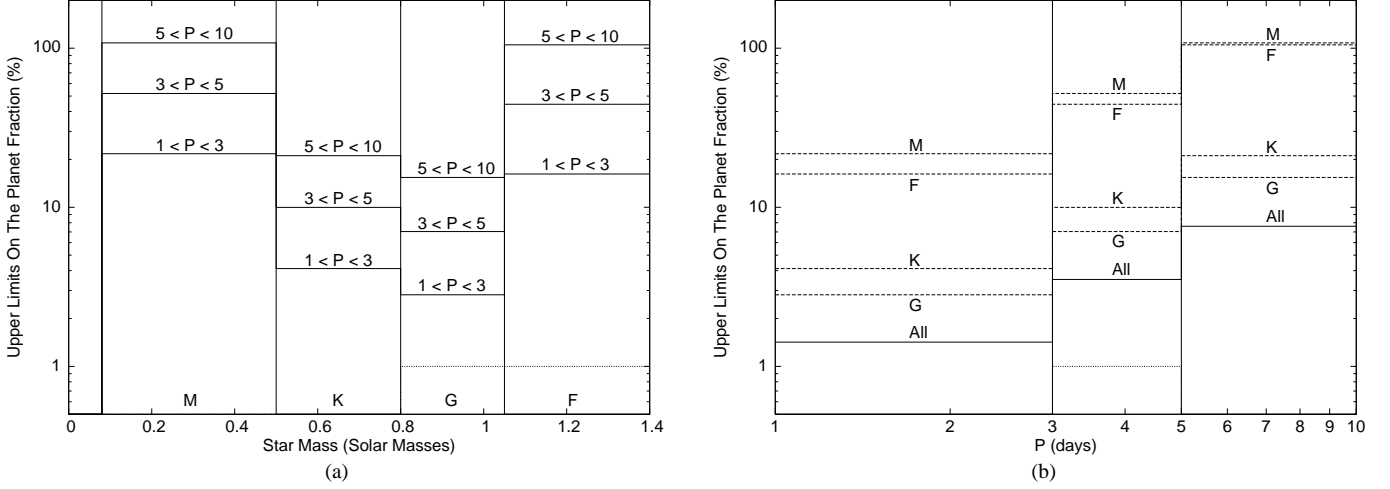


Figure 6. (a): The upper limit on f_p at $\alpha = 0.01$ as a function of star type for the various types of hot Jupiters defined by their period ranges with $R_p = 1.2R_J$. The dotted line shows the estimate of the 3-5 d hot Jupiter fraction for Solar neighbourhood Sun-like stars from Butler et al. (2000). (b): The upper limit on f_p at $\alpha = 0.01$ as a function of planet type (defined by the period range) with $R_p = 1.2R_J$ for various star types. The dotted line shows the estimate of the 3-5 d hot Jupiter fraction for Solar neighbourhood Sun-like stars from Butler et al. (2000).

also show the estimate by Butler et al. (2000) that $\sim 1\%$ of nearby Sun-like stars (late F and G dwarfs) host a 3-5 d hot Jupiter, as derived from radial velocity observations (dotted line).

It is interesting to note that although the K stars are the most numerous in our star sample (15381 stars), it is the 7423 G stars that produce the largest expected number of transiting planets, and therefore the strictest upper limits on f_p (Table 2). This is due to the fact that the G stars are generally brighter than the K stars in our sample, and therefore the corresponding gain in accuracy of the photometric measurements outweighs the smaller number of stars for which we can search for transits and the smaller transit signal for a given planetary radius.

We may compare our results directly with those of Butler et al. (2000) by considering the late F and G stars in our sample and the corresponding expected number of transiting planets for the 3-5 d hot Jupiters with $R_p = 1.2R_J$. We expected to detect ~ 10.4 such planets around the late F stars in our sample, and ~ 65.4 such planets around the G stars in our sample, an expected total of ~ 75.7 3-5 d hot Jupiters. This places an upper limit on f_p of $\sim 6.08\%$ at the 1% significance level for these types of star and planet. This is consistent with the estimate of the hot Jupiter fraction derived by Butler et al. (2000) of $f_p \approx 1\%$ and demonstrates with confidence that the hot Jupiter fraction for Sun-like stars in this field may not be more than a factor of ~ 6 times greater than that for the Solar neighbourhood.

By considering the number of planets detected by the radial-velocity technique, and by the transit technique from the OGLE survey, Gaudi, Seager & Mallén-Ornelas (2005) estimate that the relative frequency of 1-3 d very hot Jupiters to 3-9 d hot Jupiters is $0.18^{+0.12}_{-0.08}$. Using this result together with $f_p \approx 1\%$ for hot Jupiters, the authors calculate that $f_p \approx 0.1\% - 0.2\%$ for 1-3 d very hot Jupiters. This is consistent with our upper limit on f_p at $\alpha = 0.01$ of 1.42% for 1-3 d very hot Jupiters of radius $1.2R_J$.

6 CONCLUSIONS

The calculation of the expected number of transiting planet detections for a transit survey has been discussed in varying levels of detail by several authors (e.g. Gilliland et al. 2000; Weldrake et al. 2005; Hidas et al. 2005; Hood et al. 2005; Mochejska et al. 2005). A requisite for such a calculation is an estimate of the masses and radii of the stars in the sample from observational constraints or a model for the star population that predicts these properties.

In this paper, we present a method for calculating in detail the detection probabilities (and false alarm rates) of transiting planets for photometric time-series data as functions of detection threshold, planetary radius and orbital period. The calculation is based on Monte Carlo simulations and requires the properties of the host stars to be known, either from observational constraints as in the case of stellar clusters or a model for the star population. We have shown how to convert these probabilities into an expected number of transiting planet detections as a function of star and planet type. For a null result in a transit survey, this information can be used to determine a significant upper limit on the planet fraction.

In the case of the transit survey of NGC 7789 presented in BRA05 we have derived upper limits on the planet fraction for the F, G, K and M stars in the sample and for three relevant period ranges of hot Jupiters. We have also derived how these limits scale with detection threshold and planetary radius. In BRA05, it is estimated that the survey expects to detect ~ 2 HD 209458b-like transiting planets or ~ 4 OGLE-TR-56b-like transiting planets using simple arguments. Our results indicate that for HD 209458b (3-5 d hot Jupiter with $R_p \sim 1.4R_J$), and under the assumption that $f_p \approx 1\%$, we also expect to detect ~ 2 such transiting planets. Similarly for OGLE-TR-56b (1-3 d very hot Jupiter with $R_p \sim 1.2R_J$), and again under the assumption that $f_p \approx 1\%$, we expect to detect ~ 3 such transiting planets. It is encouraging to note the agreement between the two methods although the simpler method from BRA05 may tend to slightly over estimate the planetary detection rate. We conclude that the transit survey presented in BRA05 reached the sensitivity required in order to detect a few hot

Jupiters if the abundance of such planets in the field of NGC 7789 is similar to that of the Solar neighbourhood.

Improved survey design, mainly by employing a longer survey duration, will greatly improve the sensitivity to hot Jupiters. It is well known that metal rich stars have a much higher probability of hosting an extra-solar planet (Santos, Israelian, & Mayor 2004) and hence careful choice of the target star population will increase the probability of a detection. Even in the presence of a null result, a more sensitive survey will allow the derivation of tighter limits on the planet abundance.

ACKNOWLEDGEMENTS

The calculations presented in this work were done using IDL programs and the CONDOR distributed computing software. IDL is provided, under license, by Research Systems Inc. DMB is grateful to PPARC for the provision of a PhD studentship whilst at St. Andrews University and post-doctoral support at LJM, the latter as part of the “RoboNet-1.0” project.

REFERENCES

- Alonso R. et al., 2004, *ApJ*, 613, L153
 Bouchy F. et al., 2005, *astro-ph/0510119*
 Bramich D.M., 2005, *Transiting Extra-Solar Planets In The Field Of Open Cluster NGC 7789*, Ph.D. Thesis
 Bramich D.M. et al., 2005, *MNRAS*, 359, 1096
 Brown T.M., Charbonneau D., Gilliland R.L., Noyes R.W., Burrows A., 2001, *ApJ*, 552, 699
 Brown T.M., Libbrecht K.G. & Charbonneau D., 2002, *PASP*, 114, 826
 Brown T.M., 2003, *ApJ*, 593, L125
 Burrows A., Hubeny I., Hubbard W.B., Sudarsky D., Fortney J.J., 2004, *ApJ*, 610, L53
 Butler R. et al., 2000, in *Planetary Systems In The Universe: Observation, Formation And Evolution*, 1
 Charbonneau D., Brown T.M., Latham D.W., Mayor M., 2000, *ApJ*, 529, L45
 Charbonneau D., Brown T.M., Noyes R.W., Gilliland R.L., 2002, *ApJ*, 568, 377
 Drake A.J., Cook K.H., 2004, *ApJ*, 604, 379
 Gilliland R.L. et al., 2000, *ApJ*, 545, L47
 Gaudi B.S., Seager S. & Mallén-Ornelas G., 2005, *ApJ*, 623, 472
 Heacox W., 1999, *ApJ*, 526, 928
 Hidas M.G. et al., 2005, *MNRAS*, 360, 703
 Hood B. et al., 2005, *MNRAS*, 360, 791
 Lang K.R., 1992, *Astrophysical Data: Planets and Stars*. Springer, Berlin
 Mochejska B.J. et al., 2005, *AJ*, 129, 2856
 Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P., 1992, *Numerical Recipes in C. The Art of Scientific Computing*, 2nd edn. Cambridge Univ. Press, Cambridge
 Robin A.C., Reylé C., Derrière S., Picaud S., 2003, *A&A*, 409, 523
 Sato B. et al., 2005, *astro-ph/0507009*
 Santos, N. Israelian G., & Mayor M., 2004, *A&A*, 415, 1153
 Street R. et al., 2003, *MNRAS*, 340, 1287
 Udalski A. et al., 2002a, *Acta Astron.*, 52, 1
 Szewczyk O., Wyrzykowski L., Pietrzynski G., 2002b, *Acta Astron.*, 52, 115
 Udalski A., Pietrzynski G., Szymanski M., Kubiak M., Zebrun K., Soszynski I., Szewczyk O., Wyrzykowski L., 2003, *Acta Astron.*, 53, 133
 Udalski A., Szymanski M.K., Kubiak M., Pietrzynski G., Soszynski I., Zebrun K., Szewczyk O., Wyrzykowski L., 2004, *Acta Astron.*, 54, 313
 Vidal-Madjar A., Lecavelier des Etangs A., Désert J.-M., Ballester G.E., Ferlet R., Hébrard G., Mayor M., 2003, *Nature*, 422, 143
 Weldrake D.T.F., Sackett P.D., Bridges T.J., Freeman K.C., 2005, *ApJ*, 620, 1043